

CSV 合併去重工具

CSV Merge & Dedupe

合併多個 CSV / TSV 檔案，按指定欄位去重，統一編碼與欄位順序，輸出合併報告。

版本 Edition	企業授權 Enterprise
腳本 Script	csv_merge_dedupe.py
依賴 Dependencies	無（僅標準庫）
Python	3.9 或以上
貨號 SKU	SW-T-002-E
版本日期 Version	2026-10

目錄 Contents

1	功能概覽 Overview	3
2	安裝與環境 Installation	4
3	快速開始 Quick start	5
4	參數說明 Parameters	6
5	常見問題 FAQ	7
6	授權與支援 Licence & Support	8
	版權與授權聲明 Copyright & Licence	9

1 功能概覽 Overview

合併多個 CSV / TSV 檔案，按指定欄位去重，統一編碼與欄位順序，輸出合併報告。

Merge many CSV/TSV files, de-duplicate on key columns, normalise encoding and column order, and print a merge report.

- 自動偵測分隔符與 UTF-8 / GBK / Big5 編碼
- 按一個或多個欄位去重，保留首筆或未筆
- 缺少欄位自動補空，欄位順序可指定
- 輸出每個來源的筆數與去重統計

2 安裝與環境 Installation

1. 安裝 Python 3.9 或以上版本（python.org 或系統套件管理器），確認 `python --version` 可用。
2. 把 ZIP 解壓到工作目錄，保留 README、腳本與 requirements.txt。
3. 本工具只用 Python 標準庫，無需安裝第三方套件。
4. 執行 `python csv_merge_dedupe.py --help` 查看全部參數。

3 快速開始 Quick start

以下命令可直接複製到終端執行，路徑與參數請按實際情況替換。

命令示例

```
python csv_merge_dedupe.py data/*.csv -o merged.csv --key email  
python csv_merge_dedupe.py a.csv b.tsv -o out.csv --key order_id --keep last
```

4 參數說明 Parameters

參數	說明	預設
inputs	輸入檔案，支援萬用字元	—
-o / --output	—	—
--key	去重欄位，逗號分隔；不填則整行比對	—
--keep	—	first
--columns	輸出欄位順序，逗號分隔；不填則按出現順序	—
--case-insensitive	去重時忽略大小寫	—

5 常見問題 FAQ

問題	處理
提示找不到模組 (ModuleNotFoundError)	確認已在同一個 Python 環境執行 <code>pip install -r requirements.txt</code> ；建議使用虛擬環境 (<code>python -m venv .venv</code>)。
中文檔名或內容出現亂碼	用 UTF-8 保存輸入檔；Windows 終端可先執行 <code>chcp 65001</code> 。
權限不足或檔案被佔用	關閉正在使用該檔案的程式；Windows 用管理員終端，Linux/macOS 檢查檔案權限。
想先看效果不想真的改檔案	多數子命令提供 <code>--dry-run</code> 或預設乾跑，確認無誤後再加 <code>--apply</code> 。
需要定時執行	Windows 用「工作排程器」，Linux/macOS 用 <code>cron</code> ，呼叫 <code>python</code> 腳本並寫入日誌。

6 授權與支援 Licence & Support

同一集團不限使用者人數；可修改並整合進內部系統；1 年免費更新、優先電郵支援與 1 次（2 小時）線上導入協助。

Unlimited users within one corporate group; may be modified and integrated into internal systems; one year of updates, priority email support and one 2-hour online onboarding session.

- 允許：在授權範圍內運行、修改原始碼供內部使用、部署到自有伺服器。
- 不允許：轉售、公開發佈原始碼或修改版、把本工具作為付費產品的主要功能向第三方提供。
- 更新：購買後 1 年內的更新版本可透過帳戶中心重新下載；技術支援電郵 info@storywave.hk，請附貨號 SW-T-002-E。

版權與授權聲明 Copyright & Licence

本軟件工具由香港風起映海數字科技有限公司（HONG KONG FENGQI STORYWAVE DIGITAL TECHNOLOGY LIMITED）編寫並出版，受香港特別行政區及相關司法管轄區版權法律保護。

- 授權範圍：同一集團內不限人數使用；版本：企業授權；貨號：SW-T-002-E。
- 允許：在授權範圍內閱讀、列印、存入內部知識庫、在日常工作與內部培訓中使用與引用。
- 不允許：轉售、公開發佈、上傳至公開網絡、作為付費課程向第三方提供，或去除本聲明與頁眉頁腳。
- 內容為一般性方法與實務經驗分享，不構成法律、財務、醫療或其他專業意見；涉及第三方工具的功能與價格以其官方資料為準。

Published by HONG KONG FENGQI STORYWAVE DIGITAL TECHNOLOGY LIMITED. Licensed, not sold. Use within the licence scope above; no resale, public redistribution or removal of this notice. Content is general guidance and not professional advice.

聯絡 Contact

香港風起映海數字科技有限公司

香港尖沙咀梳士巴利道3號星光行7樓721A室

info@storywave.hk · +852 6658 8222 · <https://www.storywave.hk>

商業登記號碼 BR No. 81060540-000-08-26-4